



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/01, 15/11, 15/12, 15/33, 15/62, 15/70, C07K 14/00, 14/005, 14/435, 19/00, A61K 38/16, 38/17, C12Q 1/02, 1/68, 1/70	A1	(11) International Publication Number: WO 98/54311 (43) International Publication Date: 3 December 1998 (03.12.98)
(21) International Application Number: PCT/US98/10801 (22) International Filing Date: 27 May 1998 (27.05.98) (30) Priority Data: 08/863,813 27 May 1997 (27.05.97) US (63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/863,813 (CON) Filed on 27 May 1997 (27.05.97) (71) Applicant (for all designated States except US): THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10550 North Torrey Pines Road, La Jolla, CA 92037 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): BARBAS, Carlos, F., III [US/US]; 7081 Walter Street, San Diego, CA 92122 (US). GOTTESFELD, Joel, M. [US/US]; 14269 Mango Drive, San Diego, CA 92122 (US). WRIGHT, Peter, E. [US/US]; 7721 Rue Michael, La Jolla, CA 92037 (US).	(74) Agent: HAILE, Lisa, A.; Fish & Richardson P.C., Suite 1400, 4225 Executive Square, La Jolla, CA 92037 (US). (81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR		
(57) Abstract		
<p>Zinc finger proteins of the Cys₂His₂ type represent a class of malleable DNA binding proteins which may be selected to bind diverse sequences. Typically, zinc finger proteins containing three zinc finger domains, like the murine transcription factor Zif268 and the human transcription factor Sp1, bind nine contiguous base pairs (bp). To create a class of proteins which would be generally applicable to target unique sites within complex genomes, the present invention provides a polypeptide linker that fuses two three-finger proteins. Two six-fingered proteins were created and demonstrated to bind 18 contiguous bp of DNA in a sequence specific fashion. Expression of these proteins as fusions to activation or repression domains allows transcription to be specifically up or down modulated within cells. Polydactyl zinc finger proteins are broadly applicable as genome-specific transcriptional switches in gene therapy strategies and the development of novel transgenic plants and animals. Such proteins are useful for inhibiting, activating or enhancing gene expression from a zinc finger-nucleotide binding motif containing promoter or other transcriptional control element, as well as a structural gene or RNA sequence.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR

BACKGROUND OF THE INVENTION

1. *Field of the Invention*

This invention relates generally to the field of regulation of gene expression and specifically to methods of modulating gene expression by utilizing polypeptides derived from zinc finger-nucleotide binding proteins.

2. *Description of Related Art*

Transcriptional regulation is primarily achieved by the sequence-specific binding of proteins to DNA and RNA. Of the known protein motifs involved in the sequence specific recognition of DNA, the zinc finger protein is unique in its modular nature. To date, zinc finger proteins have been identified which contain between 2 and 37 modules. More than two hundred proteins, many of them transcription factors, have been shown to possess zinc fingers domains. Zinc fingers connect transcription factors to their target genes mainly by binding to specific sequences of DNA base pairs - the "rungs" in the DNA "ladder".

Zinc finger modules are approximately 30 amino acid-long motifs found in a wide variety of transcription regulatory proteins in eukaryotic organisms. As the name implies, this nucleic acid binding protein domain is folded around a zinc ion. The zinc finger domain was first recognized in the transcription factor TFIIIA from *Xenopus* oocytes (Miller, *et al.*, *EMBO*, 4:1609-1614, 1985; Brown, *et al.*, *FEBS Lett.*, 186:271-274, 1985). This protein consists of nine imperfect repeats of a consensus sequence:

(Tyr, Phe)-X-Cys-X_{2,4}-Cys-X₃-Phe-X₅-Leu-X₂-His-X_{1,4}-His-X_{2,6} (SEQ ID NO: 1)

where X is any amino acid.

- 2 -

Like TFIIIA, most zinc finger proteins have conserved cysteine and histidine residues that tetrahedrally-coordinate the single zinc atom in each finger domain. The structure of individual zinc finger peptides of this type (containing two cysteines and two histidines) such as those found in the yeast protein ADR1, the human male associated
5 protein ZFY, the HIV enhancer protein and the *Xenopus* protein Xfin have been solved by high resolution NMR methods (Kochoyan, *et al.*, *Biochemistry*, 30:3371-3386, 1991; Omichinski, *et al.*, *Biochemistry*, 29:9324-9334, 1990; Lee, *et al.*, *Science*, 245:635-637, 1989) and detailed models for the interaction of zinc fingers and DNA have been proposed (Berg, 1988; Berg, 1990; Churchill, *et al.*, 1990). Moreover, the structure of
10 a three finger polypeptide-DNA complex derived from the mouse immediate early protein zif268 (also known as Krox-24) has been solved by x-ray crystallography (Pavletich and Pabo, *Science*, 252:809-817, 1991). Each finger contains an antiparallel β -turn, a finger tip region and a short amphipathic α -helix which, in the case of zif268 zinc fingers, binds in the major groove of DNA. In addition, the conserved hydrophobic
15 amino acids and zinc coordination by the cysteine and histidine residues stabilize the structure of the individual finger domain.

While the prototype zinc finger protein TFIIIA contains an array of nine zinc fingers which binds a 43 bp sequence within the 5S RNA genes, regulatory proteins of the zif268 class (Krox-20, Sp1, for example) contain only three zinc fingers within a much larger
20 polypeptide. The three zinc fingers of zif268 each recognize a 3 bp subsite within a 9 bp recognition sequence. Most of the DNA contacts made by zif268 are with phosphates and with guanine residues on one DNA strand in the major groove of the DNA helix. In contrast, the mechanism of TFIIIA binding to DNA is more complex. The amino-terminal 3 zinc fingers recognize a 13 bp sequence and bind in the major groove. Similar
25 to zif268, these fingers also make guanine contacts primarily on one strand of the DNA. Unlike the zif268 class of proteins, zinc fingers 4 and 6 of TFIIIA each bind either in or across the minor groove, bringing fingers 5 and 7 through 9 back into contact with the major groove (Clemens, *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:10822-10826, 1992).

- 3 -

The crystal structure of zif268, indicates that specific histidine (non-zinc coordinating his residues) and arginine residues on the surface of the α -helix participate in DNA recognition. Specifically, the charged amino acids immediately preceding the α -helix and at helix positions 2, 3, and 6 (immediately preceding the conserved histidine) participate in hydrogen bonding to DNA guanines. Similar to finger 2 of the regulatory protein Krox-20 and fingers 1 and 3 of Sp1, finger 2 of TFIIIA contains histidine and arginine residues at these DNA contact positions; further, each of these zinc fingers minimally recognizes the sequence GGG. Finger swap experiments between transcription factor Sp1 and Krox-20 have confirmed the 3-bp zinc finger recognition code for this class of finger proteins (Nardelli, *et al.*, *Nature*, 349:175-178, 1989). Mutagenesis experiments have also shown the importance of these amino acids in specifying DNA recognition. It would be desirable to ascertain a simple code which specifies zinc finger-nucleotide recognition. If such a code could be deciphered, then zinc finger polypeptides might be designed to bind any chosen DNA sequence. The complex of such a polypeptide and its recognition sequence might be utilized to modulate (up or down) the transcriptional activity of the gene containing this sequence.

Zinc finger proteins have also been reported which bind to RNA. Clemens, *et al.*, (*Science*, 260:530, 1993) found that fingers 4 to 7 of TFIIIA contribute 95% of the free energy of TFIIIA binding to 5S rRNA, whereas fingers 1 to 3 make a similar contribution in binding the promoter of the 5S gene. Comparison of the two known 5S RNA binding proteins, TFIIIA and p43, reveals few homologies other than the consensus zinc ligands (C and H), hydrophobic amino acids and a threonine-tryptophan-threonine triplet motif in finger 6.

In order to redesign zinc fingers, new selective strategies must be developed and additional information on the structural basis of sequence-specific nucleotide recognition is required. Current protein engineering efforts utilize design strategies based on sequence and/or structural analogy. While such a strategy may be sufficient for the transfer of motifs, it limits the ability to produce novel nucleotide binding motifs not